FY11 Plan for **Grid & Cloud Computing/Grid Services Activity**

Prepared by: Gabriele Garzoglio, Andrew Baranovski, Philippe Canal, Dave Dykstra, Ted Hesselroth, Burt Holzman, Keith Chadwick, Tanya Levshina

Date: Apr 12, 2010

Relevant Strategic Plans - Strategic Plan for Grids, Strategic Plan for Scientific Facilities, Computing Division Strategic Plan (2010 – 2012)

Grid Services Goal

- Provide leadership in the area of middleware development for Fermilab and the Open Science Grid (OSG)
- Provide a middleware infrastructure for Fermilab and the OSG, with focus on interoperations with major peer grids, such as Enabling Grids for E-sciencE (EGEE), TeraGrid, etc., supporting the needs of Fermilab's scientific community
- Participate in and foster further externally funded activities, in line with the goals above.

Executive Summary of Objectives for FY11

Maintenance/Compliance Drivers

1. Maintenance and support of VOMRS and support for the migration to VOMS-Admin
2. Maintenance and support of Gratia (OSG-funded)
3. Maintenance and support of glideinWMS (CMS-funded)
4. Maintenance and support for the authorization infrastructure, as part of the glideinWMS project (CD and CMS-funded)
5. Maintenance and support of GIP (CMS-funded)
6. Maintenance and support of SAZ

Upgrades and Enhancement Drivers

1. Development of Gratia extensions (OSG-funded)
2. Development for the CEDPS program of a Software as a Service approach to Job and Data handling (SciDAC2-funded; ends in June 2011)
3. Development of glideinWMS (CD and CMS-funded)
4. Development and integration of GldieIn WMS and Corral into Corral WMS (STCI-funded)
5. Development of SAZ
6. Evaluate Lustre on the WAN for the CMS and Atlas applications as a contribution to the ExTENCI project (ExTENCI-funded)

Strategic Drivers

1. Development of MCAS
2. User community support for the adoption of Grid Services

Discretionary Project Drivers

1. Security reviews of software middleware
2. Modernization of the group best practices in software management

Activities and Work Definition

Activity = Grid & Cloud Computing/Grid Services/Authorization/Maintenance and Consultation

- Activity type: Service
- Description: Maintenance and consultation for Authorization and VO Registration
- Timescale: Continuous through FY11
- Metrics: Number of VOMRS releases in response to bug reports. Number of VO migrating to VOMS-Admin. Number of VO and sites interested in evaluating SVOPME for policy publication and verification.

Activity = Grid & Cloud Computing / Grid Services / Authorization / SAZ / Development

- Activity type: Project
- Description: Development of the Site AuthoriZation Service
- Timescale: Continuous through FY11
- Milestones:

Milestone	Target Completion
Integrate OSG Banning tool requirements	4th Q 2010
Functionality and Performance Testing of SAZ v3.0	4th Q 2010
External SAZ Security Review	2nd Q 2011

Activity = Grid & Cloud Computing / Grid Services / Authorization / SAZ / Management

- Activity type: Project
- Description: Coordination of the Site AuthoriZation Service project
- Timescale: Continuous through FY11
- Milestones:

Milestone	Target Completion
Project meetings	Quarterly

Activity = Grid & Cloud Computing / Grid Services / Authorization / SAZ / Support

- Activity type: Service
- Description: Support for the Site AuthoriZation Service
- Timescale: Continuous through FY11
- Metrics: Number of users banned through SAZ. Number of trouble tickets for SAZ. Number of recommended security improvements implemented.

Activity = Grid & Cloud Computing/ Grid Services / WMS / GlideinWMS / Development

- Activity type: Project
- Description: Development of the glideinWMS software
- Timescale: Continuous through FY11
- Milestones:

Milestone	Target Completion
V2.5 release	Q2 FY11
V3.0 release	Q3 FY11
V3.1 release	Q4 FY11

Activity = Grid & Cloud Computing / Grid Services / WMS / GlideinWMS / Maintenance and Support

- Activity type: Service
- Description: Maintenance of code and user support
- Timescale: Continuous through FY11
- Metrics: Number of trouble tickets reported and resolved for the system. Turn-around time in resolving incidents. Number of stakeholders.

Activity = Grid & Cloud Computing/Grid Services / WMS / GlideinWMS / Management and Outreach

- Activity type: Project
- Description: Project management and outreach to new potential stakeholders
- Timescale: Continuous through FY11
- Milestones: Quarterly stakeholders meeting

Activity = Grid & Cloud Computing / Grid Services / WMS / GlideinWMS / Corral

- Activity type: Project
- Description: Integration of the glideinWMS system with Corral as per funded NSF grant.
- Timescale: Continuous through FY11
- Milestones:

Milestone	Target Completion
v0.1 Release (initial CorralWMS integration)	Q1 FY11
v0.2 Release (multicore CorralWMS)	Q2 FY11
v0.3 Release (monitoring prototype)	Q3 FY11

Activity = Grid & Cloud Computing / Grid Services / Information Systems

- Activity type: Service
- Description: Maintenance and support of the Grid information system (GIP/ BDII)
- Timescale: Continuous through FY11
- Metrics: Number of trouble tickets reported and resolved for the system. Turn-around time in resolving incidents.

Activity = Grid & Cloud Computing / Grid Services / Accounting / Maintenance

- Activity type: Service
- Description: Software maintenance and improved quality assurance processes of the Gratia accounting system, including US CMS use cases and an integration of CD accounting use cases.
- Timescale: Continuous through FY11
- Metrics: Number of issues about non-working or inaccurate report/data. Turn-around time to resolve these issues.

Activity = Grid & Cloud Computing / Grid Services / Accounting / Management

- Activity type: Project
- Description: Coordination of the activities for the Gratia accounting system
- Timescale: Continuous through FY11
- Milestones:

Milestone	Target Completion
Stakeholders' meeting	Weekly

Activity = Grid & Cloud Computing/Grid Services/Metrics Management/MCAS/Development

- Activity type: Project
- Description: Development activities for the Metrics Correlation and Analysis Service
- Timescale: Continuous through FY11
- Milestones:

Milestone	Target Completion
Web based data visualization tool kit (Minos)	3th Q 2010
Dashboard description API and its implementation to map workflows into JBoss and vanilla UI containers	4th Q 2010
Define and implement dCache metric collection and analysis Workflows	1th Q 2011
Implement dCache monitoring dashboards	2th Q 2011
Implement automated workflow and Warehouse stress test infrastructure	2th Q 2011
Develop metric analysis wizards for common applications (dCache,Condor,PhedEx)	4th Q 2011
Project documentation	Continuous

Activity = Grid & Cloud Computing/Grid Services/Metrics Management/MCAS/Management and Outreach

- Activity type: Project
- Description: MCAS management and engagement of new stakeholders
- Timescale: Continuous through FY11
- Milestones:

Milestone	Target Completion

Weekly project meetings	Weekly
Stakeholders engagement	Monthly
Establish and further develop collaboration with BNL	Monthly

Activity = Grid & Cloud Computing/Grid Services/CEDPS/Storage

- Activity type: Project
- Description: Investigate, improve, and troubleshoot storage solutions
- Timescale: Continuous through FY11 – Official project end date is June 2011
- Milestones:

Milestone	Target Completion
Design and negotiate an approach for resilient handling of GlideinWMS process output data using globus.org service	2th Q 2010
Develop Phase I prototype. Implementation of resilience layer.	3th Q 2010
Develop Phase I pre-production service	4th Q 2010
Develop Phase II pre-production service. Implementation of long term failover as part of the resilience layer.	1th Q 2011
Project closure documents	2nd Q 2011

Activity = Grid & Cloud Computing/Grid Services/CEDPS/MCAS

- Activity type: Project
- Description: Participate to troubleshooting activities of CEDPS via relevant activities in the MCAS project
- Timescale: Continuous through FY11 – Official project end date is June 2011
- Milestones:

Milestone	Target Completion
Define and implement template based transformation codes for time based correlation of time stamped XML data	4th Q 2010

Activity = Grid & Cloud Computing / Grid Services / Security

- Activity type: Service
- Description: Security reviews for the Grid Services software and increased group acumen.
- Timescale: Continuous through FY11
- Metrics: Number of security reviews performed. Evaluation of security acumen at security-related meetings and discussions.

Activity = Grid & Cloud Computing / Grid Services / Infrastructure Modernization

- Activity type: Project
- Description: Modernization of the group best practices in software management
- Timescale: Continuous through FY11
- Milestones:

Milestone	Target Completion
Establish regularly attended developer's forum	4th Q 2010
Determine testing standards for software projects.	1st Q 2011
Establish testing standards in target software projects.	4th Q 2011

Activity = Grid & Cloud Computing / Grid Services / User Support

- Activity type: Service
- Description: Assist the Grid user community in the adoption of Grid Services and related best practices.
- Timescale: Continuous through FY11
- Metrics: Number of communities properly adopting Grid Services.
-

Activity = Grid & Cloud Computing / Grid Services / ExTENCI

- Activity type: Project
- Description: Evaluate Lustre on the WAN for the CMS and Atlas applications as a contribution to the ExTENCI project.
- Timescale: Funds might end in May FY11 or might be extended 1 more year.
- Milestones:

Milestone	Target Completion
Deploy WAN Lustre system	4th Q 2010
Evaluate WAN Lustre for CMS application	2nd Q 2011

Detailed Tactical Plan Objectives and Priorities

Maintenance and Compliance Drivers

Objectives:

1. Maintain the infrastructure for VO membership registration. Support the OSG VOs and the FermiGrid administrators, who operate VOMRS for OSG, to migrate to VOMS-admin. Investigate new mechanisms for VO and site policy definition, publication, and enforcement.
2. Work closely with stakeholders to identify and appropriately prioritize their maintenance needs from the Gratia software stack (text and graphical reports, probes and collectors).
3. Provide maintenance and support for the glideinWMS for CMS, Fermigrid, CDF, OSG, and other stakeholders.
4. In the context of glideinWMS, package gLexec authorization software in collaboration with gLite developers along with OSG-specific components.

Provide maintenance and support to OSG.

5. Contribute to the maintenance and support of the OSG grid information system (GIP).
6. Develop and maintain the SAZ service to enable user/vo/role/ca banning on campus grid facilities, in particular on FermiGrid, and to provide support to customers of the SAZ software.

Assumptions and Risks (may be 2 sections if it makes your points clearer)

1. Failure to transition to VOMS-Admin as a registration service will incur in the need for continuous maintenance of VOMRS, estimated at up to 0.1 FTE.
2. Failure to understand stakeholders' maintenance needs for the Gratia accounting system will limit the ability of introducing new metrics to measure the functional properties of the OSG as a system.
3. Failure to operate and support glideinWMS may have a major effect on the efficiency of ongoing data analyses for CMS, CDF, and the Minos experiment.
4. Failure to package new gLexec releases for OSG may prohibit wide-spread use of pilot-based workload management systems (glideinWMS, ATLAS's PANDA) on OSG, lowering efficiency and possibly decreasing the pool of potentially available resources.
5. Failure to maintain and support the OSG information system would seriously impact the efficiency of CMS analysis on OSG, disrupt data transfer for the US LHC Tier 1 sites, and negatively affect computing work for other OSG organizations.
6. Failure to maintain and support the SAZ service will result in increased operational complexity in reacting to security incidents for the FermiGrid facility.

Upgrades and Enhancement Drivers

Objectives:

1. Ensure that (potential) Gratia Extensions provided by external projects are coordinated and well integrated into the existing code base, test, release and support mechanisms. Improve quality assurance processes for the software.
2. In the context of the CEDPS and GlideinWMS programs, provide reliable mechanism for storing output data. Interface globus.org as implementation of point-to-point data transfers. Enhance MCAS for the troubleshooting use cases of CEDPS.
3. As part of the CorralWMS project, develop glideinWMS so that it is fully compatible and integrated with the Corral frontend. Design and develop an application-level monitoring system. Establish collaborative workspaces, training methods, and contribute to project leadership.
4. Enhance and further develop glideinWMS based on stakeholder input.
5. Evolve the SAZ service to integrate the user-banning requirements by OSG and to improve software maintenance and operations.
6. Evaluate Lustre on the WAN for the CMS and Atlas applications as a contribution to the ExTENCI project. Measure Lustre's performance, observe its

robustness and ease of use, and determine whether it can fully support the applications at small remote (Tier-3) sites. Report on it to the larger community. Contribute administrative knowledge of Lustre in the tuning of system parameters. Demonstrate ability to collaborate with TeraGrid counterparts toward a common goal of sharing a national cyber-infrastructure.

Assumptions and Risks:

1. Failure to appropriately coordinate external effort to expand Gratia will result in probable fork of the project. This and the lack of improved quality assurance processes will result in additional support effort by customer and stakeholder and possibly compatibility issues.
2. According to CMS, 50% of data failures happen while handling process output data. Increasing failure resilience of the stage-out phase of the standalone process is difficult without leveraging features of both GlideinWMS and data movement services such as globus.org. CEDPS at Fermilab depends on and focuses those technologies.
3. Failure to meet the CorralWMS objectives would have a negative effect on the spread of pilot-based workload management systems and increase the cost for organizations wishing to use both OSG and TeraGrid. As the project is externally funded, failure to make progress in FY11 could also impact the funding of the effort.
4. Failure to respond to new stakeholder requests may delay deployment by new and potential stakeholders, and decrease the deployment level by existing stakeholders, effectively lowering efficiency across the OSG cyberinfrastructure.
5. As one of the recognized Policy Decision Points on the Grid, SAZ has integrated the XACML-based module to achieve authorization interoperability on the Grid. This module has the potential to reduce the maintenance needs of SAZ through software reuse. The ability to interoperate and reduce maintenance cost, however, is at risk as the XACML-based module might need performance improvements before if can be universally deployed at Fermilab and on the OSG.
6. Without participating in the ExTENCI project, we will likely fail to understand if Lustre is a valuable alternative to dCache and Hadoop as a storage system for the CMS and Altlas applications. In addition, the ExTENCI project is an opportunity for Fermilab as a member of the OSG community to demonstrate our ability to collaborate with our TeraGrid counterparts towards a single national cyber-infrastructure. Failure to participate in the project might negatively impact the perception of the funding agencies in our capabilities as collaborators.

Strategic Drivers

Objectives:

1. Based on the existing MCAS infrastructure, develop high level automations to implement ‘canned’ use cases of metrics collection, preparation, and display in the following applications areas: dCache, Minos, and the neutrino experiments.

2. Assist the Grid user community in the adoption of Grid Services and related best practices. The targeted community will include academic institutions, virtual organizations related to Fermilab, and campus grids.

Assumptions and Risks:

1. The two key features of the MCAS project are
    a. providing quick and easy access to building custom user-level dynamic reports
    b. documenting the schema of user/system metrics

   Lack of common services to generate reports on metrics will encourage users to do uncoordinated development/integration to achieve objectives similar to the ones of MCAS. These activities will inevitably lead to duplication of effort and potential software vulnerabilities, due to the lack of common development standards. Failure to formalize and catalog metrics will hinder the ability to report the performance of computing services operations to stakeholders, including to CD management.
2. Failure to assist the user community with the adoption of Grid Services will increase the diversity of the solutions for already addressed problems and possibly increase the load to support Grid Services.

Discretionary Projects Drivers

Objectives:

1. Perform security-focused reviews of several software projects.
2. Analyze emerging software management practices, such as agile software methodologies, and understand their adoption for the Grid Services development group. Modernize the software management practices and tools. Improve on software quality assurance techniques. Organize a forum and periodic roundtables for CD software developers.

Assumptions and Risks:

1. Many production level software products perform security related functions as part of their normal operations. Insuring that these products conform to security practices decreases the possibility of security incidents.
2. Failure to modernize our software management practices will result in loss of efficiency in the management of software development processes. It may also result in the disenfranchising of the software development workforce.

Staffing Issues:

Gratia development and user community support activities (including OSG Engagement) need the hiring of a new developer, as per an already approved requisition. Staffing for FY10 was lower than anticipated by ~0.4 FTE due to the temporary reassignment of Eileen Berman from the Grid Department Head to the Service Desk.